

Academic Research

Designing a Formative Measure for Corporate Reputation

Sabrina Helm

Heinrich-Heine-University of Duesseldorf, Germany

ABSTRACT

The construct of reputation has attracted much attention among researchers and practitioners alike. The growing body of literature in this area has led to a wide variety of measurement approaches, albeit most publications do not point out the epistemic nature of reputation as a construct. It is not clear (1) if reputation is a formative or reflective construct and (2) how it should be conceptualized using a formative approach. The incorrect specification of the construct would produce misleading results, which in turn would provide no solid basis for reputation management. Based on a study among German consumers, this paper conceptualizes reputation as a formative construct consisting of ten indicators. The findings suggest that this conceptualization is a solid method on which to build reputation measures, which in turn are an important step for efficient reputation management.

KEYWORDS: *formative indicators, reflective indicators, reputation measurement, reputation scales, structural equation modelling*

INTRODUCTION

There has been a resurgence of interest in the reputation construct among researchers and practitioners (Fombrun, 2003; Dowling, 2001). The growing body of literature has led to an abundance of different definitions of corporate reputation. Gotsi and Wilson (2001) for instance highlight numerous viewpoints (Mahon, 2002; Wartick, 2002). They conclude that corporate

reputation is 'a stakeholder's overall evaluation of a company over time' (Gotsi and Wilson, 2001: 29).

From an empirical point of view, corporate reputation has remained a rather controversial issue. For instance, Nguyen and Leblanc (2001) diagnose a lack of consensus on valid scales for reputation; they therefore propose to use direct measures. Unfortunately, this does not lead to practical insights for reputation management because the sources of a good or bad reputation do not become evident. Effective measures are still considered the biggest hurdles for reputation management (Larkin, 2003).

In addition to the lack of consensus on the dimensions of reputation, recent publications challenge the common approach of specifying complex constructs such as reputation in general (Rossiter, 2002; Jarvis *et al.*, 2003). Given the developments increasing the usability of *structural equation modelling* in the social sciences (Bollen, 1989; Baumgartner and Homburg, 1996), the relevance of this discussion for the conceptualization of reputation needs to be investigated in more detail.

In structural equation modelling, every construct or latent variable is assigned a set of indicators, but in publications, the epistemic relationship between variable and indicators is often not considered. Latent variables may be associated with *reflective* or *formative* indicators. Most researchers assume a reflective relationship, meaning

that the unobserved latent variable effects the indicators. In this case, all indicators ‘measure the same thing and should covary at a high level if they are good measures of the underlying variable’ (Bagozzi, 1994: 331). If the latent variable is thought to be a construct of all of its indicators — like an index or ranking — it needs to be measured formatively. ‘Formative indicators give rise to the unobserved theoretical construct. In this case the empirical indicators produce or contribute to the construct’ (Fornell, 1982: 8).

As Hulland (1999) claims, it is very important from a conceptual and methodological standpoint, which kind of indicator specification is used. Diamantopoulos and Siguaw (2002: 11) emphasize that the ‘alternative approaches to deriving measures can produce substantially different operationalization of the same construct’.

The present paper contributes to the emerging literature on reputation by focusing on the construct’s epistemic structure. More specifically, it raises two research questions:

- 1 How do reflective and formative indicator specifications for corporate reputation differ?
- 2 How can a formative approach to measuring reputation be developed?

In order to address the research questions, the paper is structured as follows: In the next section, the relevant literature on reputation measures is reviewed and the differences between a reflective and a formative measurement approach are discussed. A four-step approach to formative indicator specification is followed in section 3. In section 4, the empirical research design of the study will be described. Section 5 discusses major research findings. Structural equation modelling is used for further data analysis and validation. Finally, section 6 presents limitations of the present study and suggests directions for future research.

REPUTATION: A REFLECTIVE OR A FORMATIVE CONSTRUCT?

Measures of Reputation

The growing interest in reputation has led to the development of a variety of different construct measures. Among others, Fombrun (1998), Lewis (2001) and Wartick (2002) have reviewed the existing measurement approaches, highlighting *Fortune*’s annual ‘Most Admired Companies’ and the Reputation Institute’s ‘Reputation Quotient’SM (RQ) as the most frequently used and discussed data sets. Both represent rankings of companies based on a cluster of different corporate associations that represent different stakeholders’ expectations regarding the activities of a company (Berens and van Riel, 2004). Examples for such social expectations are the delivery of high-quality products, treating employees fairly, and delivering a good financial performance. Indicators used to measure corporate reputation usually represent one facet of these expectations. The set of indicators is then aggregated to make up the construct of reputation in the sense of an index.

Besides this measurement approach, Berens and van Riel (2004) identify the concept of corporate personality. This means that corporate reputation is measured using indicators that represent personality traits of people and which may also be attributed to organizations. The measure developed by Davies *et al.* (2003; 2004) is a prominent example of this conceptualization of reputation. Furthermore, trust-based reputation measures are named as a third dominant approach. According to Berens and van Riel (2004), the corporate credibility scale developed by Newell and Goldsmith (2001) represents this method. Here, the perception of an organization’s honesty, reliability and benevolence are predictors of corporate behavior and therefore interpreted as possible indica-

tors of reputation within measurement models.

Besides these main measurement approaches, Berens and van Riel (2004) identify some less prominent ones. Also, Bromley (2002), van Riel *et al.* (1998), as well as Thevissen (2002) describe different methods for measuring corporate reputation which are not necessarily based on quantitative, large-scale surveys. As the first mentioned, expectations-based approach may be classified as the one used and discussed most frequently in the literature and the one best known in practice, the paper will focus on this particular approach.

The Epistemic Nature of Reputation

In spite of numerous research projects on measuring reputation, publications on the specification of reputation in regard to its epistemic nature are quite rare. Generally, rankings or indexes are classical examples of formative construct conceptualization (Bagozzi, 1994; Diamantopoulos and Siguaw, 2002), and rankings are the most common method of measuring reputation today (Fombrun, 1998; Fombrun *et al.*, 2000; Larkin, 2003). The process describing the generation of the underlying items is usually not published. Also, applications in the context of structural equation models (SEMs) are rare. By using such models, the relationship between reputation and other constructs, for instance stakeholder satisfaction or corporate profitability, may be investigated. Reputation within an SEM-context is discussed by Yoon *et al.* (1993), Andreassen and Lindestad (1998) and Nguyen and Leblanc (2001). Incorrect specification of the constructs might lead to inaccurate conclusions about the structural relationships between them (Jarvis *et al.*, 2003) and to misleading managerial implications.

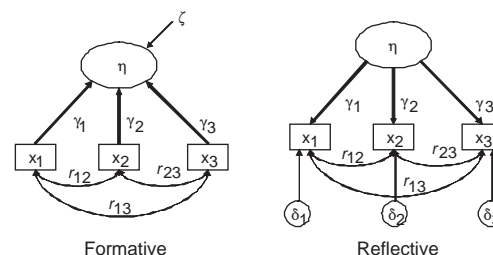
SEMs contain two inter-related models. The *measurement model* defines the constructs or latent variables and assigns

observed variables — the indicators — to them. The *structural model* defines the causal relationship between the latent variables or constructs (Bollen, 1989; Gefen *et al.*, 2000). The question of formative or reflective indicators is an issue when specifying the measurement model. It refers to the links between the construct and its indicators that represent the *epistemic relationship* or 'rules of correspondence' (Hulland, 1999; Bagozzi, 1994). These are shown in Figure 1.

Using *reflective indicators*, the researcher assumes that the observable indicators are reflections or representations of the construct, which means that the construct should be unidimensional and the items correlated. An increase in one indicator is accompanied by increases of the other indicators (Chin and Newsted, 1999; Jarvis *et al.*, 2003). During the scale purification process, items with low loadings are usually taken out of the construct (Hulland, 1999; Gefen *et al.*, 2000).

Formative indicators 'cause' the latent variable, they represent different dimensions of it. The construct is then a summation of the observed variables with which it is associated. The indicators do not have to be correlated or to represent the same underlying dimension (Bollen and Lennox, 1991; Gefen *et al.*, 2000). An increase in

Figure 1: Epistemic relationships between the latent variable and its indicators



Source: Diamantopoulos and Winklhofer, 2001: 270

one indicator does not have to be accompanied by increases of the other indicators (Chin and Newsted, 1999).

Looking at the construct of reputation, explicit statements regarding its epistemic nature have not been dealt with in detail in the literature. Andreassen (1994) uses a reflective approach to measure reputation and employs six indicators ranging from 'good services' to 'being inventive' and 'management competence'. Caruana and Chircop (2000) develop a scale for reputation by following the common guidelines for reflective measures such as developing a set of interchangeable items tapping each aspect of the construct, eliminating items due to low inter-item correlation and thereby improving coefficient alpha. Although the authors do not explain the epistemic nature of their construct, it can be concluded that they presume it to be reflective. The same might be said for Fombrun and Shanley (1990), Cordeiro and Schwalbach (1999) and Yoon *et al.* (1993), who develop or criticize reputation models and report the loadings of the items used and the correlations between the reputation attributes. These reports are needed for reflective constructs, but not for formative ones. Finally, in their list of incorrectly specified constructs, Diamantopoulos and Winklhofer (2001) present the reflective construct 'company reputation' developed by Goldberg and Hartwick (1990). It is measured as a firm's reputation with employees, financial investors, the US public and the Canadian public, and it should be formative.

For several years now, the Reputation Institute has worked on a refined measurement approach for reputation (Fombrun *et al.*, 2000). The epistemic background of the RQ is however, not mentioned in publications. As the RQ is designed to be 'a valid, reliable and robust tool for measuring corporate reputation' (Gardberg and Fombrun, 2002: 306), this shortcoming needs to

be resolved. The construct itself is obviously formative: It consists of 20 indicators assigned to six dimensions (not factors). Combined, the dimensions represent the measure of reputation (Gardberg and Fombrun, 2002), which is typical for a formative construct. The process followed to identify the dimensions and indicators is typical for reflective measures. For instance, loadings of the indicators on the dimensions are reported. Formative indicators are reported by their respective weight (Seltin and Kees, 1994; Fornell and Cha, 1994). Also, items were included in the model and eliminated afterwards due to insufficient loadings. This procedure is again typical for reflective measures as the latent variable is represented in every indicator. Furthermore, multi-item measures are used for each dimension. Taking into account the complex structure of the RQ, it can be hypothesized that in its creation, reputation was thought of as a second-order, multidimensional construct. This means that the RQ-dimensions, such as emotional appeal or social responsibility are formative ingredients forming the construct of reputation, whereas the indicators used to measure each dimension are conceptualized reflectively. Some of the affluent RQ-indicators might be redundant if the measure were systematically conceptualized as a formative construct.

How do the two kinds of specification of the reputation construct differ? If for instance reputation were to be understood as 'the aggregation of a single stakeholder's perceptions of how well organizational responses are meeting the demands and expectations of many organizational stakeholders' (Wartick, 1992: 34) and conceptualized as a construct consisting of perceptions on product quality, treatment of employees, management quality, care for the environment etc., the relationship between these different aspects of reputation (the indicators) and the construct

could indeed be twofold.

If reputation in this case were understood as a reflective construct, this would mean that the indicators — such as product quality, management quality, treatment of employees — are ‘effects of a construct’ (Bollen and Lennox, 1991: 305). Reputation would lead to these effects meaning that *reputation determines* the quality of products, the quality of management, the treatment of employees, and so forth. These effects would be alternative outcomes of reputation.

Conceptualizing reputation as a formative construct would mean that the indicators lead to the construct as *input*. Reputation would be an aggregation of all its indicators such as product quality or treatment of employees. This would imply that because it delivers high quality products, a firm has a good reputation; because it treats employees right, it has a good reputation and so forth.

The reflective approach is appropriate for explaining the effects of reputation, rather than its formation as a construct. Because of its positive reputation, a firm has to keep on producing high quality products in order not to damage its reputation. ‘To be held in high esteem creates obligations that managers and companies must live up to’ (Fombrun, 1996: 10). Hence, the *effectiveness of reputation management* (understood as a separate construct) should be conceptualized reflectively. The construct reputation itself has a *formative nature* if it is measured based on the social expectations approach as discussed by Berens and van Riel (2004) and uses a firm’s performance attributes such as product quality or treatment of employees as sub-dimensions of reputation. This is also illustrated by the fact that, eg an increase in product quality must not necessarily be accompanied by increases in the other performance domains such as treatment of employees, management qualification etc., which would be

descriptive of reflective indicators (Chin and Newsted, 1999). The same argument applies for the measure developed by Goldberg and Hartwick (1990) mentioned above. If reputation is measured as a collection of indicators referring to the perception of the firm’s reputation by different stakeholder groups, leaving out any of those indicators (i.e. stakeholder groups) would change the construct.

In the following sections, the process and results of conceptualizing reputation as a formative construct are discussed on the basis of an empirical study.

SCALE DEVELOPMENT

Within the context of this paper, reputation as perceived by the company’s stakeholders is to be measured using formative indicators. The process of developing such a measurement model is different from the process of building scales for reflective latent variables (Baumgartner and Homburg, 1996). Diamantopoulos and Winklhofer (2001) suggest a *four step approach* for index construction that includes content and indicator specification, indicator collinearity and external validity, which was implemented in our study of reputation.

Content Specification

The first step includes content specification, that is the *definition of the construct*. Usually, a wide understanding of the formative construct is called for because ‘failure to consider all facets of the construct will lead to an exclusion of relevant indicators’ (Diamantopoulos and Winklhofer, 2001: 271). On the other hand, ex-post deletion of ‘weak’ indicators is usually not considered an option (Bagozzi, 1994; Nunnally and Bernstein, 1994). Drawing on existing scales and the literature on reputation, two focus group interviews and 40 individual interviews with consumers, employees and shareholders were conducted prior to our study to ensure that the definition of repu-

tation would capture its domain of content sufficiently. These qualitative measures aim at granting *content validity* of the construct — the kind of validity Rossiter (2002) claims to be most essential when defining measurement models.

In the understanding of the interviewees, reputation of a firm is not a single person's perception of a firm or his/her attitude towards that firm. Reputation is a *collective construct* meaning that a person will base his/her evaluation of a firm on what is said about that firm (by the public) and on what his/her perceptions are of the firm's capability and willingness to fulfill stakeholders' needs. This is for instance supported by Emler (1990: 181) who claims: 'reputations are social, not individual judgments'. Still, individuals are the source of information about a firm's reputation as Wartick (2002: 374) clarifies: 'reputation, be it corporate or otherwise, cannot be argued to be anything but purely perceptual'. As one respondent expressed, reputation is 'not what I personally think about a company or how I experienced their service, but it is about what everyone says and thinks about that company'. Taking a look at the literature on reputation, a wide range of definitions becomes obvious but no consensus on its operationalization emerges. Most survey approaches seem to interpret reputation as the aggregation of stakeholders' attitudes towards a firm — without specifying the difference between those stakeholders' own experiences and attitudes, and information on the firm based on 'hearsay' (Fombrun *et al.*, 2000; Gardberg, 2001). These individual perceptions are then aggregated, forming what Bromley (2002: 36) calls a 'meta reputation' — a fusion of a large collection of personal judgments about a standard set of corporate attributes'. According to our qualitative interviews, reputation is a perceptual construct but is based on perceptions of what others might think or actually say

about a firm. As another interview partner said: 'What I think of a company and what their reputation is like, that might be two different things'. Therefore, *corporate reputation* is defined as a single stakeholder's perception of the estimation in which a certain firm is held by its stakeholders in general. This definition comes very close to Fombrun's (1996: 3) understanding of reputation as 'the overall estimation in which a company is held by its constituents', but it closes the gap between perceptions and collective assessments.

Indicator Specification

The second step is the *generation of the formative indicators* making up the construct. All facets of the construct need to be covered by these indicators in order to grant a fit between the definition and its operationalization. Other than with reflective indicators, this set of indicators cannot 'be chosen randomly from the universe of items relating to the construct of interest' (DeVellis, 1991: 55). Also, the indicators are not interchangeable in the formative case. The whole construct has to be covered by the indicators that build up a formative measurement model (Bollen and Lennox, 1991). In the case of reputation, all facets of the estimation in which a firm is held by stakeholders have to be included, calling for a variety of indicators capturing the different contributions of the firm to different stakeholders' needs. But in contrast to the reflective case, it is not necessary for every facet of the construct to be measured by multiple indicators (Churchill, 1979; Bagozzi and Baumgartner, 1994; Rossiter, 2002).

The approach used to create the scale relies upon the idea that the same scale can be applied to a wider range of companies. It integrates focus group and individual interviews as well as a literature review. Potential indicators were first identified by looking for items commonly used in prior

measurement models of corporate reputation. As this review process lacks input by the potential respondents, interviews were conducted.

Interviewees were selected based on a convenience sample. The author's private and business contacts were used to identify possible interview partners within these contacts' social networks. Using this snowball-approach, 40 interviewees were contacted and questioned in their roles as customers, shareholders, or employees. They were to find definitions of 'reputation', 'corporate reputation' and to name the characteristics of companies with a bad or good reputation. These characteristics mainly reflect the expectations that respondents have regarding a company's actions (Berens and van Riel, 2004). Considering a possible 'stakeholder-bias' of reputation measures, it might be expected that respondents in the role of customer might name different characteristics of reputable firms than other stakeholders. But this was not the case: all three groups of interviewees used comparable terms to describe companies with a good or a bad reputation, especially emphasizing the quality of products, treatment of employees and treatment of the environment.

After this inductive step, two focus group interviews were conducted. Members were students and fellow researchers from different academic backgrounds. They were asked to discuss the definition of corporate reputation as stated above and to name and discuss characteristics of companies with a good or bad reputation, the importance of reputation for firms and stakeholders and examples of firms with a very good or very bad reputation. The interactive character of this type of interview provided numerous details about the perception of reputation. Combining this, inductive and deductive procedures in the scale development process aims at developing a measure that is applicable to a wide

scale of different companies as the literature review and the input of the interviewees were not restricted to a special industry, company, or a single stakeholder group's statements. All in all, a list of 59 possible indicators was gathered from both types of interviews. Not all of them are equally useful for construct development as they only partially tapped their intended construct, also tapped other constructs, or were not very important for the evaluation of reputation. Therefore, a three-step pretest approach was employed to refine the scales. In the first step, students and colleagues at the university were asked to check for overlap between the items suggested. For instance, nominations such as 'economic success', 'financial soundness' or 'financial health' were grouped and labeled 'financial performance'. Also, they were to determine on a 4-point scale how important the remaining 25 items are in explaining if a company has a good or bad reputation. This ensures that only indicators perceived as relevant for a company's reputation are included in the measure. Table 1 lists these items and their evaluation by the 26 respondents in columns 1 to 7.

In a second step, an item-sorting task as proposed by Anderson and Gerbing (1991) — again using 12 fellow academics — showed how well the items tapped into their constructs. Participants were told about the basic research design, the definition of reputation and two other constructs, customer satisfaction and loyalty. They were asked to assign individual items to what they believed to be the 'correct' construct out of the set of three. Two indices proposed by Anderson and Gerbing (1991) — the *proportion of substantive agreement* (p_{sa}) and the *substantive-validity coefficient* (c_{sv}) — were computed for each item to identify those that were difficult to assign to the corresponding construct. For reputation, this led to a reduced item set of 12 as only items with a p_{sa} of above 0.75

Table 1: Importance and Assignment of Items

No.	Company characteristic	N_1	1	2	3	4	N_2	n_c	n_o	p_{sa}	c_{sv}
1	Corporate success	26	12	14	0	0	12	11	1	0.92	0.83
2	Customer orientation ¹	26	16	10	0	0	12	9	3	0.75	0.50
3	Quality of products ¹	26	19	7	0	0	12	10	1	0.83	0.75
4	Respectability of activities on markets ²	26	13	13	0	0	12	10	2	0.83	0.66
5	Familiarity of company brands	26	14	8	4	0	12	6	4	0.50	0.17
6	Sincerity concerning the information of the public ²	26	12	14	0	0	12	11	1	0.92	0.83
7	Market leadership	26	12	10	2	2	12	7	5	0.58	0.17
8	Innovation potential	26	9	11	6	0	12	6	4	0.50	0.17
9	Stability of market presence	26	10	14	2	0	12	8	4	0.67	0.33
10	Value for money of products ¹	26	11	14	1	0	12	9	2	0.75	0.58
11	Superiority to competitors	26	9	12	5	0	12	7	4	0.58	0.25
12	Service offers for customers	26	6	13	7	0	12	4	6	0.33	-0.17
13	Credibility of advertising claims	26	7	15	4	0	12	10	1	0.83	0.75
14	Commitment to protecting the environment	26	9	13	4	0	12	12	0	1	1
15	Consideration of consumer rights	26	6	12	8	0	12	8	2	0.67	0.50
16	Investment in advertising/ frequency of advertising	26	5	10	6	5	12	7	3	0.58	0.33
17	Continuity in advertising	26	4	9	9	4	12	6	5	0.50	0.08
18	Corporate philosophy ³	26	4	8	9	5	12	9	2	0.75	0.58
19	Taking responsibility for public matters ³	26	3	9	11	3	12	10	1	0.83	0.75
20	Person of CEO	26	4	10	8	4	12	6	3	0.50	0.25
21	Commitment to charitable causes (eg social or cultural causes)	26	5	15	6	0	12	11	1	0.92	0.83
22	Qualification of Management	26	4	17	5	0	12	10	2	0.83	0.66
23	Treatment of employees	26	5	15	6	0	12	12	0	1	1
24	Financial performance of firm	26	2	22	2	0	12	9	3	0.75	0.50
25	Attractiveness as investment	26	0	12	9	5	12	6	4	0.50	0.17

N_1 : Number of respondents for pretest on importance of characteristics

Scale (columns 4–7), Characteristics named in column 2 are 1 = very important; 2 = important; 3 = not so important; 4 = not at all important for the estimation in which a certain company is held

N_2 : Number of respondents for pretest on item assignment to constructs

p_{sa} : Proportion of substantive agreement, calculated as n_c/N_2 (number of ‘correct’ assignments divided by the number of respondents)

c_{sv} : Substantive-validity coefficient, calculated as $(n_c - n_o)/N_2$ (number of ‘correct’ assignments minus number of most frequent assignment divided by the number of respondents)

¹The assignment test also included items for the constructs ‘customer satisfaction’ and ‘customer loyalty’. The former was assigned items that were very much similar to some reputation items, but were labeled as ‘own experience with product quality’, ‘own experience with customer orientation’, etc. This explains why items 2, 3 and 10 were assigned to the construct ‘corporate reputation’ by quite a number of respondents in spite of their obvious closeness to the satisfaction construct

²Item was eliminated after the final pretest step as respondents did not interpret the terms respectability and sincerity homogeneously concerning firms and their market relationships or were unsure about their meanings

³In spite of a sufficient result in the item-sorting test, the item was left out of the scale. It was not voted to be very important in the formation of reputation by a sufficient amount of respondents (threshold-value: at least 60%/16 of the respondents rated the item as ‘very important’ or ‘important’)

and a c_{sv} above 0.5 were kept in the measure (see also Table 1; items kept are marked in grey).

In a third step, questionnaires were administered to 20 consumers in a ‘think aloud’ answer mode. This led to the elimination of two more items and to minor revisions on wording. The ten remaining reputation indicators were then included in the final survey (see Table 2). In accordance with Rossiter (2002: 322), bipolar entirely verbalized 7-point scales were used.

Multicollinearity

The *third step* in the development process is to check for indicator collinearity. Multicollinearity is a relevant problem for formative indicators as the formative measurement model is based on a multiple regression. Sample size and the strength of indicator correlations affect the stability of the indicator coefficients (Diamantopoulos and Winklhofer, 2001). Therefore, a high amount of collinearity between indicators makes it problematic to separate distinct influences of the individual indicators on

the latent variable. The variance inflation factor should not exceed ten (Kleinbaum *et al.*, 1998; Hair *et al.*, 1995).

External Construct Validity

The *fourth and last step* in construct specification is checking external construct validity. An internal consistency perspective is not appropriate for formative measures due to their very nature (Diamantopoulos and Winklhofer, 2001). Spector (1992) suggests correlating each indicator to another latent variable (external to the main construct). Such an external criterion could be ‘a global item that summarizes the essence of the construct that the index purports to measure’ (Diamantopoulos and Winklhofer, 2001: 272). A second approach to ‘qualify’ formative indicators for the measurement model is to include the entire construct in a wider nomological context, meaning that other constructs and their relationships to the construct in question have to be measured (Bagozzi, 1994). If the construct has the theoretically hypothesized impact on the other constructs in the structural model, this confirms the nomological

Table 2: Information on the Formative Measurement Model for Reputation

Indicator	Description	Weight	t-value
x ₁	Quality of products	0.2733	5.1129
x ₂	Commitment to protecting the environment	0.2204	3.1633
x ₃	Corporate success	0.0183	0.3245
x ₄	Treatment of employees	-0.0243	0.2915
x ₅	Customer orientation	0.1287	2.7766
x ₆	Commitment to charitable and social issues	0.0987	1.4297
x ₇	Value for money of products	0.3029	6.1554
x ₈	Financial performance	0.0218	0.3988
x ₉	Qualification of management	0.0308	0.5454
x ₁₀	Credibility of advertising claims	0.3023	5.7450

Evaluations were based on the question: ‘Concerning the following characteristics, does company x have a good or bad reputation in the public?’, measured on a 7-point scale: 1 = ‘a very good reputation’, 4 = ‘a neutral reputation, neither good nor bad’, 7 = ‘a very bad reputation’.

As PLS does not directly provide standard errors the bootstrap procedure of PLS-Graph was used to calculate the t-statistics (Tenenhaus, 2003).

Threshold t-values: $t = 1.64$ for $p = 0.05$; $t = 2.33$ for $p = 0.01$.

validity of the measurement models used (Diamantopoulos and Winklhofer, 2001). For identification purposes, it is important to include one or several reflective latent variables in the structural model (Bollen, 1989).

RESEARCH DESIGN AND SAMPLE STRUCTURE

The Empirical Study

The procedure and results described below were part of a larger research project that was focused at measuring the reputation of an international consumer goods producer from the viewpoint of three of his stakeholder groups. Here, the results of the study conducted in the German consumer sample will be discussed, whereas the results of the other surveys on shareholders and employees are not subject of this paper.

Data Collection

The respondents were contacted for a personal interview conducted by a leading research institute. A total of 1,681 consumers aged between 20 and 60; 80 per cent female (main target group of the consumer goods producer), were contacted at their homes following a random-route design. Interviews took place at 210 sample points all across Germany and were computer-aided, which allowed for item rotation (Dillman, 2000; Rossiter, 2002).

In 729 cases, the household or targeted person refused to take part in the interview, leading to a response rate of 56.6 per cent (952 cases), which is within the usual range of personal interviews of consumers (Dillman, 2000). As consumers cannot always provide valid answers concerning a company's reputation (Carlson, 1963; Berens and van Riel, 2004), two questions were used to filter respondents who (a) knew the company name and were customers of the company's products and (b)

felt knowledgeable about its reputation. This reduced the effective sample size to 762 usable questionnaires.

RESULTS

Due to the problems in specifying formative indicators with software for covariance structure analysis such as LISREL (MacCallum and Browne, 1993), partial least squares (PLS) was used for data analysis; the software package used was PLS-Graph. Covariance-based methods such as LISREL require very large data sets when formative measurement models are involved. 'The frequency of improper and uninterpretable solutions advise against the use of LISREL unless its assumptions are verifiably true and its objectives consistent with the objectives of the study; and, when they are not, PLS presents a viable alternative' (Fornell and Bookstein, 1982: 313). PLS allows for incorporating formative as well as reflective measurement models and is distribution-free (Jöreskog and Wold, 1982; for an overview and a discussion of other features, see Fornell and Bookstein, 1982; Fornell and Cha, 1994; Wold 1982. For a comparison of LISREL and PLS, see Chin and Newsted, 1999; Gefen *et al.*, 2000).

An evaluation of formative measurement models is based on the weights of the indicators. 'The weights provide information as to what the make-up and relative importance are for each indicator in the creation/formation of the component' (Chin, 1998: 307). The results for the reputation construct are shown in Table 2, with weights that are not significant at $p=0.05$ marked in grey.

Five indicators have a weight below 0.1; one has a negative sign. Seltin and Keeves (1994) claim such indicators to be 'trivial' and call for their removal in order to build parsimonious models (Jöreskog and Wold, 1982). Concerning formative variables, indicator deletion is problematic though as 'omitting an indicator is omitting a part of

the construct' (Bollen and Lennox, 1991: 308). Facets of the reputation construct would be removed resulting in the formation of a new construct. Rossiter (2002: 315) therefore claims: 'Item selection to increase the "reliability" of the formed scale is definitely not appropriate'. In our case, reputation would be reduced to product quality, environmental issues, customer orientation, credibility of advertising claims and value for money. These might well be the most important facets of reputation from a consumer's point of view, but if the aim of the researcher is to build a reputation measure applicable to different stakeholder groups (as is also the stated objective of the Reputation QuotientSM; Fombrun, 1998; Fombrun *et al.*, 2000), an elimination of items would reduce the capacity of the measurement model to cover other stakeholder groups as well.

The literature gives no specific answer to the question of how indicator elimination should be handled in the formative case (Diamantopoulos and Winklhofer, 2001; Bollen and Lennox, 1991). In their model, Fornell *et al.* (1990: 252) also find some very weak weights, including one with a negative sign. They conclude: '*in this particular sample*, this variable did not add to our initial conceptualization of this theoretical construct', which is not sufficient to prove that the model is not valid. They therefore decided not to delete any indicators. Diamantopoulos and Winklhofer (2001: 273) caution: 'a non-significant *t*-statistic for γ [= weight of the indicator; author's note] fails to reject the zero value hypothesis'. They suggest eliminating non-significant indicators starting with the one with the lowest *t*-value, but only as long as the breadth of construct definition is not reduced by this elimination procedure. 'How to balance these considerations is a question that has not yet been fully resolved. Indicator elimination — by whatever means — should not be divorced

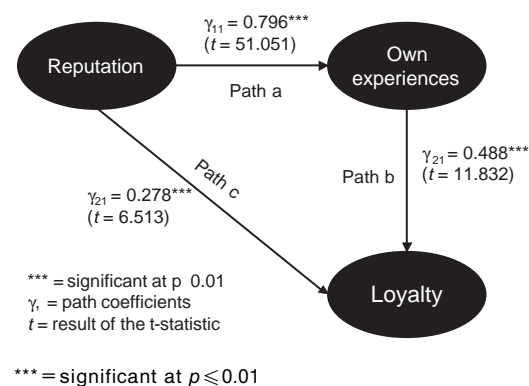
from conceptual considerations when a formative measurement model is involved' (Diamantopoulos and Winklhofer, 2001: 272). In order to illustrate the methodology, the whole set of reputation indicators is contained in the reputation measure.

In a next step, the model is to be tested regarding multicollinearity and external construct validity. The highest value of the variance inflation factor was 3.09 for the reputation construct, so multicollinearity should not pose a problem as it is below the common cut-off threshold of ten (Kleinbaum *et al.*, 1998).

In order to check for external construct validity, each indicator is to be correlated to an external criterion that summarizes the construct to be measured (Spector, 1992; Diamantopoulos and Winklhofer, 2001). Therefore, an additional question was integrated in the study measuring the overall reputation of the consumer goods producer as perceived by the respondents. All reputation indicators correlate positively and significantly (at $p < 0.05$ or better) with this global reputation measure (highest coefficient: 0.59 for 'quality of products'; lowest coefficient: 0.38 for 'treatment of employees').

A second approach for evaluating external construct validity is to include the measure in a wider nomological network. The construct of reputation was part of the

Figure 2: Information on the structural model



SEM-model shown in Figure 2. As mentioned, the study of reputation described here was part of a larger research project, which aimed at exploring the relationship between reputation, a stakeholder's personal experience with the consumer goods producer (as an important part of his satisfaction) and his loyalty towards this company. Here, this model is only used to examine the nomological validity of the reputation construct. Therefore, only the measures relevant for our methodological investigation of formative measures will be discussed, omitting the theoretical foundations of the suggested relationships between the constructs. The hypothesized relationships were a direct, positive impact of reputation on the respondent's evaluation of his own experience, a direct impact of experience on loyalty and a direct impact of reputation on loyalty. Consumers' experience was modeled as a formative construct, loyalty was thought to be reflective in accordance with the literature (Andreassen, 1994; Oliver, 1997).

All path coefficients are significant at $p=0.01$. The strongest effect shows path a from reputation to the evaluation of own experiences. The direct path c from reputation to loyalty is rather weak. All paths are significant, meaning that external construct validity of the measurement model for reputation is affirmed (Diamantopoulos and Winklhofer, 2001).

DISCUSSION AND LIMITATIONS OF THE STUDY

The growing body of knowledge on reputation has been dominated by research focused on the construct itself; inclusion of reputation in structural equation modelling in order to examine its interplay with other constructs has remained scarce. Also, the epistemic nature of complex constructs such as reputation has largely been ignored in the literature. It has however been gaining interest recently (Diamantopoulos and

Winklhofer, 2001; Rossiter, 2002).

The present study investigated the epistemic nature of reputation and the attributes that form the construct. It aimed at contributing to a better knowledge of (1) the differences between formative and reflective conceptualizations of reputation and (2) how to develop a formative measure of corporate reputation.

Reputation indicators were developed based on a literature review, focus group interviews, and personal interviews. Ten reputation indicators were identified. This limited number contributes to the conciseness of the new measure and meets the criterion of parsimony (Seltin and Keeves, 1994; Hair *et al.*, 1995). Concerning construct validity, some indicators produced low weights and were not significant. Due to the formative nature of the construct however, no indicators were eliminated. The empirical research confirms the nomological validity of the construct.

As in any empirical research, the results of the present study cannot be interpreted without taking into account its limitations. Furthermore, this study generates a set of research-worthy questions that need to be addressed in future projects.

First, the study (not the generation of indicators) was restricted to the reputation of one consumer goods producer. The measurement model needs to be validated in different industry settings, for different firms within an industry and for different stakeholder groups. As the measure was developed considering different stakeholders' views about firms, it can be used as a vehicle for comparing customer perceptions of a firm's reputation with other stakeholders, thereby contributing to closing a gap in the literature (Scholes and Clutterbuck, 1998).

Secondly, the handling of low-weighted items needs to be investigated more closely. As some authors claim, different stakeholder groups — in a conceptual sense —

‘put different weight’ on reputation dimensions so that firms have multiple ‘reputations’ (Fombrun and Shanley, 1990). An application of the model in different stakeholder-settings might result in different weights of the indicators.

The question of the epistemic nature of reputation needs to be discussed concerning every measurement model of reputation. This also applies to the most prominent examples such as *Fortune*’s annual ‘Most Admired Companies’ and the Reputation Institute’s RQ. Both measures are based on a set of indicators that form an index of a company’s characteristics as perceived by respondents. Accepting that this is a formative structure implies that the process illustrated above needs to be followed when conceptualizing the measure. Consequently, the measure portrayed here might prove a (formative) alternative to the RQ and related, reflectively modeled constructs that purport to measure reputation based on the social expectations approach. Taking different approaches to measure reputation, such as the corporate personality- or trust-based conceptualizations as discussed by Berens and van Riel (2004) also merits further investigation into the epistemic structure of the measures.

In conclusion, it becomes evident that researchers need to be aware of the conceptual differences between the measurement approaches and clearly identify their reputation model’s epistemic nature. The paper’s aim was to contribute to a better understanding of the different conceptualizations which helps to build better measures of reputation. This is an important step aiming at efficient reputation management.

REFERENCES

- Anderson, J.C. and Gerbing, D.W. (1991) ‘Predicting the performance of measures in a confirmatory factor analysis with a pretest assessment of their substantive validities’, *Journal of Applied Psychology*, 76, 732–740.
- Andreassen, T.W. (1994) ‘Satisfaction, loyalty and reputation as indicators of customer orientation in the public sector’, *International Journal of Public Sector Management*, 7, 16–34.
- Andreassen, T.W. and Lindestad, B. (1998) ‘The effect of corporate image in the formation of customer loyalty’, *Journal of Service Research*, 1, 82–92.
- Bagozzi, R. (1994) ‘Structural equation models in marketing research: Basic principles’, in: R. Bagozzi (ed.) *Principles of Marketing Research*, Blackwell, Oxford, 317–385.
- Bagozzi, R. and Baumgartner, H. (1994) ‘The evaluation of structural equation models and hypothesis tests’, in R. Bagozzi (ed.) *Principles of Marketing Research*, Blackwell, Oxford, 386–422.
- Baumgartner, H. and Homburg, C. (1996) ‘Applications of structural equation modeling in marketing and consumer research: A review’, *International Journal of Research in Marketing*, 13, 139–161.
- Berens, G. and van Riel, C.B.M. (2004) ‘Corporate associations in the academic literature: Three main streams of thought in the reputation measurement literature’, *Corporate Reputation Review*, 7, 161–178.
- Bollen, K. (1989) *Structural Equations with Latent Variables*, Wiley, New York.
- Bollen, K. and Lennox, R. (1991) ‘Conventional wisdom on measurement: A structural equation perspective’, *Psychological Bulletin*, 110, 305–314.
- Bromley, D.B. (2002) ‘Comparing corporate reputations: League tables, quotients, benchmarks, or case studies?’, *Corporate Reputation Review*, 5, 35–50.
- Carlson, R.O. (1963) ‘The nature of corporate images’, in: J.W. Riley (ed.) *The Corporation and its Publics*, Wiley, New York/London, 24–47.
- Caruana, A. and Chircop, S. (2000) ‘Measuring corporate reputation: A case example’, *Corporate Reputation Review*, 3, 43–57.
- Chin, W.W. (1998) ‘The partial least squares approach to structural equation modeling’, in: G.A. Marcoulides (ed.) *Modern Methods for Business Research*, Erlbaum, Mahwah/London, 295–336.
- Chin, W.W. and Newsted, P.R. (1999) ‘Structural equation modeling analysis with small samples using partial least squares’, in: R.H. Hoyle and H. Rick (eds) *Statistical Strategies for Small Sample Research*, Sage, Thousand Oaks, 307–341.
- Churchill, G.A. (1979) ‘A paradigm for developing better measures of marketing constructs’, *Journal of Marketing Research*, 16, 64–73.
- Cordeiro, J.J. and Schwalbach, J. (2000) ‘Preliminary evidence on the structure and determinants of global corporate reputations’, *Working Paper of the Institute for Management*, Faculty of Economics, Humboldt-University, Berlin.

- Davies, G., Chun, R., da Silva, R. and Roper, S. (2004) 'A corporate character scale to assess employee and customer views of organizational reputation', *Corporate Reputation Review*, 7, 125–146.
- Davies, G., Chun, R., da Silva, R. and Roper, S. (2003) *Corporate Reputation and Competitiveness*, Routledge, London.
- Diamantopoulos, A. and Siguaw, J.A. (2002) 'Formative vs. reflective indicators in measure development: Does the choice of indicators matter?', *The Center For Hospitality Research Working Papers*, Ithaca.
- Diamantopoulos, A. and Winklhofer, H.M. (2001) 'Index construction with formative indicators: An alternative to scale development', *Journal of Marketing Research*, 38, 269–277.
- DeVellis, R.F. (1991) *Scale Development: Theory and Applications*, Sage Publications, Newbury Park.
- Dillman, D.A. (2000) *Mail and Internet Surveys*, 2nd ed., Wiley, New York.
- Dowling, G.R. (2001) *Creating Corporate Reputations*, Oxford University Press, New York.
- Emler, N. (1990) 'A social psychology of reputation', *European Review of Social Psychology*, 1, 171–193.
- Fombrun, C. (2003) *Fame and Fortune*, Prentice Hall, Upper Saddle River.
- Fombrun, C. (1998) 'Indices of corporate reputation: An analysis of media rankings and social monitors' ratings', *Corporate Reputation Review*, 1, 327–340.
- Fombrun, C. (1996) *Reputation*, Harvard Business School Press, Boston.
- Fombrun, C., Gardberg, N. and Sever, J. (2000) 'The Reputation QuotientSM: A multi-stakeholder measure of corporate reputation', *The Journal of Brand Management*, 7, 241–255.
- Fombrun, C. and Shanley, M. (1990) 'What's in a name? Reputation building and corporate strategy', *Academy of Management Journal*, 33, 233–258.
- Fornell, C. (1982) 'A second generation of multivariate analysis — An overview', in: C. Fornell (ed.) *A Second Generation of Multivariate Analysis*, Greenwood Press, New York, 1–21.
- Fornell, C. and Bookstein, F.L. (1982) 'A comparison analysis of two structural equation models: LISREL and PLS applied to market data', in: C. Fornell (ed.) *A Second Generation of Multivariate Analysis*, Greenwood Press, New York, 289–324.
- Fornell, C. and Cha, J. (1994) 'Partial least squares', in: R.P. Bagozzi (ed.) *Advanced Methods of Marketing Research*, Blackwell, Cambridge, 52–78.
- Fornell, C., Lorange, P. and Roos, J. (1990) 'The cooperative venture formation process: A latent variable structural modeling approach', *Management Science*, 36, 1246–1255.
- Gardberg, N. (2001) 'How do individuals construct corporate reputations?', Dissertation, New York University, New York.
- Gardberg, N. and Fombrun, C. (2002) 'The global reputation quotient project: First steps towards a cross-nationally valid measure of corporate reputation', *Corporate Reputation Review*, 4, 303–315.
- Gefen, D., Straub, D.W. and Boudreau, M.-C. (2000) 'Structural equation modeling and regression: Guidelines for research practice', *Communications of the Association for Information Systems*, 4.
- Goldberg, M.E. and Hartwick, J. (1990) 'The effects of advertiser reputation and extremity of advertising claim on advertising effectiveness', *Journal of Consumer Research*, 17, 172–179.
- Gotsi, M. and Wilson, A.M. (2001) 'Corporate reputation: Seeking a definition', *Corporate Communications*, 6, 24–30.
- Hair, J.F., Anderson, R.E., Tatham, R.L. and Black, W.C. (1995) *Multivariate Data Analysis*, 4th edn., Prentice Hall, Englewood Cliffs.
- Hulland, J. (1999) 'Use of partial least squares (PLS) in strategic management research: A review of four recent studies', *Strategic Management Journal*, 20, 195–204.
- Jarvis, C., Mackenzie, S.B. and Podsakoff, P.M. (2003) 'A critical review of construct indicators and measurement model misspecification in marketing and consumer research', *Journal of Consumer Research*, 30, 199–218.
- Jöreskog, K.G. and Wold, H. (1982) 'The ML and PLS technique for modeling with latent variables — Historical and comparative aspects', in: K.G. Jöreskog and H. Wold (eds), *Systems under Indirect Observation*, North Holland, Amsterdam, 263–270.
- Kleinbaum, D.G., Kupper, L., Muller, K.E. and Nizam, A. (1998) *Applied Regression Analysis and Other Multivariable Methods*, 3rd edn. Duxbury Press, Pacific Grove.
- Larkin, J. (2003) *Strategic Reputation Risk Management*, MacMillan, Basingstoke.
- Lewis, S. (2001) 'Measuring corporate reputation', *Corporate Communications*, 6, 31–35.
- MacCallum, R.C. and Browne, M.W. (1993) 'The use of causal indicators in covariance structure models: Some practical issues', *Psychological Bulletin*, 114, 533–541.
- Mahon, J.F. (2002) 'Corporate reputation', *Business & Society*, 41, 415–445.
- Newall, S.J. and Goldsmith, R.E. (2001) 'The development of a scale to measure perceived corporate

- credibility', *Journal of Business Research*, 52, 235–247.
- Nguyen, N. and Leblanc, G. (2001) 'Corporate image and corporate reputation in customers' retention decision in services', *Journal of Retailing and Consumer Services*, 8, 227–236.
- Nunnally, J.C. and Bernstein, I.H. (1994) *Psychometric Theory*, 3rd edn. McGraw-Hill, New York.
- Oliver, R.L. (1997) *Satisfaction*, McGraw-Hill, New York *et al.*
- Rossiter, J.R. (2002) 'The C-OAR-SE procedure for scale development in marketing', *International Journal of Research in Marketing*, 19, 305–335.
- Scholes, E. and Clutterbuck, D. (1998) 'Communication with stakeholders: An integrated approach', *Long Range Planning*, 31, 227–238.
- Seltn, N. and Keeves, J.P. (1994) 'Path analysis with latent variables', in: T. Husén, T. Postlethwaite (eds), *The International Encyclopedia of Education*, 2nd edn. Pergamon, Oxford, 4352–4359.
- Spector, P.E. (1992) *Summated Rating Scales Construction*, Sage, Newbury Park.
- Tenenhaus, M. (2003) 'Comparison between PLS and LISREL Approaches for Structural Equation Modeling: Application to the Measure of Customer Satisfaction', unpublished Working Paper of HEC School of Management, Jouy-en-Josas.
- Thevissen, F. (2002) 'Corporate reputation in the eye of the beholder', *Corporate Reputation Review*, 4, 318–326.
- van Riel, C.B.M., Stoeker, N.E. and Maathuis, O.J.M. (1998) 'Measuring corporate images', *Corporate Reputation Review*, 1, 313–326.
- Wartick, S.L. (2002) 'Measuring corporate reputation', *Business & Society*, 41, 371–392.
- Wartick, S.L. (1992) 'The relationship between intense media exposure and change in corporate reputation', *Business & Society*, 31, 33–49.
- Wold, H. (1982) 'Systems under indirect observation using PLS', in: C. Fornell (ed.) *A Second Generation of Multivariate Analysis*, North Holland, New York, 325–347.
- Yoon, E., Guffey, H.J. and Kijewski, V. (1993) 'The effects of information and company reputation on intentions to buy a business service', *Journal of Business Research*, 27, 215–228.

Copyright of Corporate Reputation Review is the property of Henry Stewart Publications and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.